

Dokumentation der Web-Statistik 2006 des German Microdata Lab

Hagen, Susanna; Lüttinger, Paul; Urban, Jeanine

Veröffentlichungsversion / Published Version
Arbeitspapier / working paper

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Hagen, S., Lüttinger, P., & Urban, J. (2007). *Dokumentation der Web-Statistik 2006 des German Microdata Lab*. (GESIS-Methodenberichte, 2/2007). Mannheim: GESIS-ZUMA. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-206832>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

GESIS-Methodenbericht Nr. 2/2007

**Dokumentation der Web-Statistik 2006
des German Microdata Lab**

Susanna Hagen, Paul Lüttinger, Jeanine Urban

November 2007

GESIS-ZUMA
German Microdata Lab
Postfach 12 21 55
68072 Mannheim
Tel.: +49 621 1246 265
Telefax: +49 621 1246 100
Email: paul.luettinger@gesis.org

GESIS-Methodenberichte

Die GESIS ist ein Institut der Leibniz-Gemeinschaft.

ISSN:

1865-7575

Herausgeber, Druck
und Vertrieb:

GESIS
Postfach 12 21 55, 68072 Mannheim

Printed in Germany

Gliederung

1.	Einleitung.....	3
1.1	Datenanforderung.....	3
1.2	Auswertungsstatistik	5
2.	Ergebnisse	6
2.1	Gesamtentwicklung.....	6
2.2	Daten.....	8
2.3	Service	14
2.4	Publikationen.....	19
3.	Zusammenfassung	22

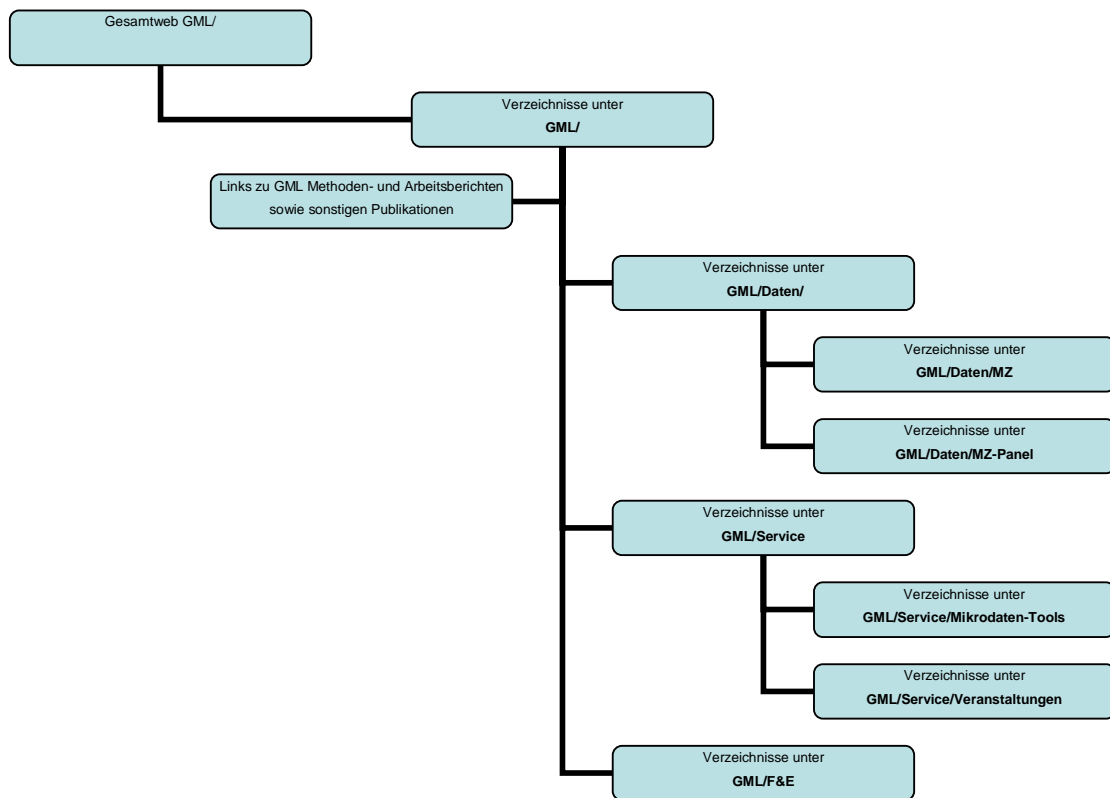
1. Einleitung

Die Webseiten des German Microdata Lab (GML) sind ein zentraler Bestandteil des Serviceangebotes. Über die Webseiten werden den Nutzern alle Metadaten zu den neu aufbereiteten Scientific Use Files der Mikrozensus zur Verfügung gestellt und die bereits vorhandenen Seiten laufend überarbeitet und aktualisiert. Dies ist z.B. dann notwendig, wenn Daten neu hinzukommen, wie die älteren Mikrozensus aus den 70er und 80er Jahren und diese Daten sich in wesentlichen Merkmalen, wie z.B. der Stichprobenziehung oder den Gewichtungungsverfahren von den neueren Daten unterscheiden. Zudem werden über die Webseiten „Werkzeuge“, so genannte „Tools“, für den Umgang mit den Daten angeboten, Workshops und Konferenzen angekündigt, Publikationen zum Download bereitgestellt und vieles mehr. Der folgende Bericht gibt einen Überblick über die Nutzung der Webseiten des GML für das Jahr 2006.

1.1 Datenanforderung

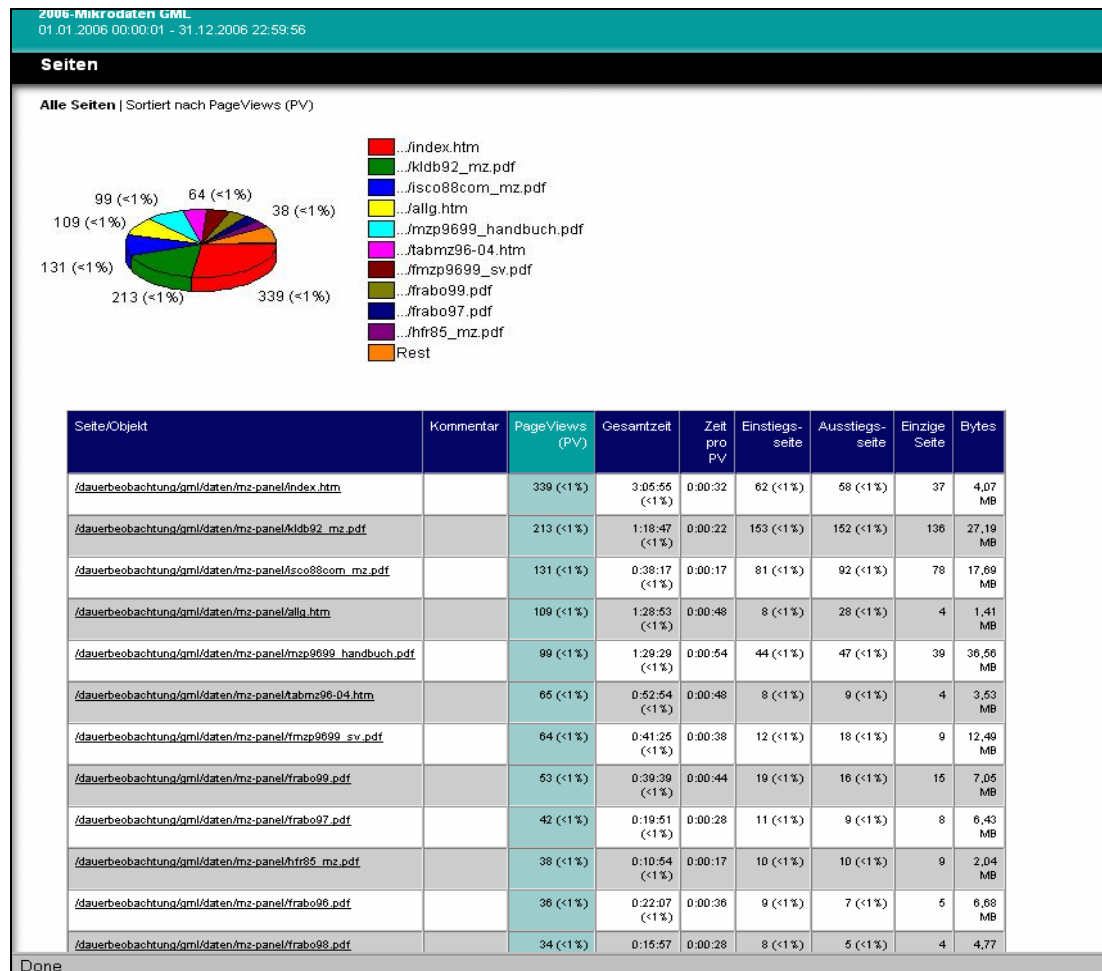
Die Web-Statistik des GML für 2006 orientiert sich an der Struktur der Webseiten, wie sie in Abbildung 1 dargestellt sind.

Abbildung 1: Struktur der durch die Web-Statistik erfassten Seiten des GML-Gesamtwebs



Diesem Aufbau folgend wurde im Detail festgelegt, welche Seiten von der Statistik erfasst werden sollen. Die Statistik selbst wurde vom GESIS-IZ (Berlin), welche den GESIS Web-Server betreut, automatisch generiert. In der Abbildung 2 ist die Zuordnung der Daten, wie sie vom GESIS-IZ geliefert werden, beispielhaft aufgeführt. Der Schwerpunkt der Web-Statistik für 2006 liegt auf den für das GML zentralen Metadatenangeboten zu den verschiedenen Scientific Use Files des Mikrozensus und der Einkommens- und Verbrauchsstichprobe sowie den Mikrodaten-Tools, die den größten Teil der Seiten des GML-Gesamtwebs ausmachen. Ein weiteres Auswertungsmerkmal war ferner die Anzahl der Downloads von PDF-Dateien der vom GML bereitgestellten Dokumente (Arbeits- und Methodenberichte, Publikationen), die einen direkten Hinweis auf die Nutzungshäufigkeit dieses Angebotes liefern.

Abbildung 2: Beispiel für eine Auswertungsstatistik



1.2 Auswertungsstatistik

Zur Auswertung wird das Analysetool WebSuxess sowie eine Anzahl von Filterskripten verwendet, mit deren Hilfe die Zugriffe auf den GESIS Webserver herausgefiltert werden, die z.B. von den verschiedenen GESIS Standorten oder von Suchmaschinen erfolgen.¹

Bei jedem Besuch eines „Clienten“ wird vom Programm WebSuxess eine Log-Datei angelegt. In dieser Log-Datei werden verschiedene Parameter gespeichert, wie z.B. das Datum des Zugriffs, die Uhrzeit, die IP-Adresse des Clienten, Anzahl der übertragenen

1) Zur Vorgehensweise bei der Auswertung von Log-Dateien und damit verbundener methodischer Probleme siehe: Mell, Wolf-Dieter 2002: Methodische Anmerkungen zur Auswertung der WWW-Log-Dateien des Servers www.gesis.org. IZ-Arbeitsbericht Nr. 26.

Bytes, Browser des Klienten, Anzahl der Hits, Visits, Page Views etc. Die wesentliche Kennziffer für die Nutzung der Seiten ist die Anzahl der so genannten Page Views (PV). Darunter wird die Anzahl der Sichtkontakte mit einer Seite verstanden, d.h. wie oft eine Seite aufgerufen wird. Anders als bei den ebenfalls oft verwendeten Kenngrößen „Hits“ oder „Visits“, die jedes einzelne Element einer Seite (Grafik, Frame) zählen, sobald eine Seite aufgerufen wird, wird bei einem Page View nur die komplette Seite erfasst, unabhängig davon, aus wievielen Elementen (Dateien, Frames) sich die Seite zusammensetzt.

Aber auch die Page Views sind als reine Kenngröße nur bedingt aussagekräftig, da das Problem der Zugriffe von so genannten Proxy-Servern nicht gelöst ist. Proxy-Server greifen stellvertretend für einen Nutzer auf eine Seite zu und lagern diese in einem Zwischenspeicher (cache) ab. Greift der Nutzer erneut auf diese Seite zu, wird die Anfrage – außer es erfolgt eine Aktualisierung dieser Seite – vom cache Speicher beantwortet und nicht mehr vom GESIS-Server direkt. Dadurch werden die Zugriffe untererfasst, so dass mit Sicherheit nur gesagt werden kann, „daß die Anzahl der Anfragen eine Mindestangabe ist und die tatsächliche Anzahl [...] augenblicklich nicht einmal geschätzt werden“ [kann].² Im Folgenden wird ausschließlich die Zahl der Page Views verwendet.

2. Ergebnisse

2.1 Gesamtentwicklung

Das Web-Angebot des GML ist in den letzten Jahren kontinuierlich um mehrere hundert Seiten pro Jahr gewachsen und umfasste zum Ende des Jahres 2006 insgesamt 5.759 Seiten (Tabelle 1). Das Web-Angebot des GML ist damit innerhalb der GESIS das umfassendste Angebot: Ein Drittel der Seiten des gesamten GESIS Webs von 15.733 Seiten werden durch das GML bereitgestellt. Innerhalb des Bereichs „Dauerbeobachtung“, das die Abteilungen GML, Soziale Indikatoren und ALLBUS/ISSP umfasst, stellt das GML 89 Prozent der vom gesamten Bereich Dauerbeobachtung angebotenen Seiten. Die Zahl der Page Views für das GML ist von 2003 bis 2005 kontinuierlich angestiegen. 2006 erfolgte ein deutlicher Rückgang, allerdings auf einem weiterhin sehr hohen Niveau. Diese Verringerung in der Anzahl der Page Views resultiert aus einer

2) http://193.175.239.64/statistik/GESIS_deutsch/GESIS_MA/info.html

technischen Änderung: Während bislang die Ausprägungen von Variablen (pro Erhebungsjahr etwa 300) auf jeweils einzelnen Seiten dargestellt wurden, erfolgt dies seit der Bereitstellung des Mikrozensus 2004 im Jahr 2006 effizienter auf nur vier Seiten.

Im Jahr 2006 erfolgten täglich über 1.500 Page Views auf eine der Webseiten des GML (Tabelle 1). Damit entfallen 10 Prozent aller Page Views für das gesamte GESIS Webangebot auf Seiten des GML.

**Tabelle 1: Webseiten und Zugriffe
(GML-Angebot: Deutsche und englische Seiten)**

	Anzahl der Seiten	Page Views	PV pro Tag
GML 2003	3.594	490.791	1.344
GML 2004	4.415	716.106	1.961
GML 2005*	5.046	850.299	2.329
GML 2006	5.759	550.549	1.508
Dauerbeobachtung 2006	6.474	1.111.052	3.043
GESIS Gesamt 2006	15.733	5.133.392	14.064

* bis April 2005: /dauerbeobachtung/mikrodaten und /dauerbeobachtung/einkommen
ab Mai 2005: dauerbeobachtung/gml/

Eine weitere Kenngröße für die generelle Nutzungshäufigkeit von Web-Seiten ist die Anzahl der gesendeten Bytes, die beim Aufruf einer Seite an den Rechner des Nutzers geschickt werden (Tabelle 2). Seit 2001 zeigt sich eine deutliche positive Entwicklung in der Nutzung der Webseiten der GESIS und des GML. Insbesondere im Verlaufe des Jahres 2007 ergab sich nochmals eine deutliche Steigerung gegenüber 2005 bzw. 2006 auf 44 Gigabyte, was mehr als eine Verdoppelung im Volumen der gesendeten Gigabytes gegenüber den vorhergehenden Jahren bedeutet.³

Tabelle 2: Anzahl der gesendeten Gigabytes, 2001-2007

	2001	2002	2003	2004	2005	2006	2007
GESIS Gesamt	59	109	135	213	341	503	563
davon: Dauerbeobachtung	18	25	41	62	97	69	143
GML	7	9	12	20	21	18	44
GML, % von Gesis Gesamt	11,9	8,2	8,8	9,4	6,2	3,6	7,8

3) Auch diese Zahlen sind nur Mindestangaben, da die Seiten, die von einem cache Speicher abgerufen werden, nicht mehr als Download erfasst werden.

2.2 Daten

Der Großteil der Webseiten des GML ist den Informationen über Daten gewidmet. Im Jahr 2006 enthalten 89 Prozent der Seiten (Tabelle 3, Spalte 2) und 86 Prozent der Kontakte zu den Webseiten des GML erfolgen in diesem Bereich (Tabelle 3, Spalte 4).

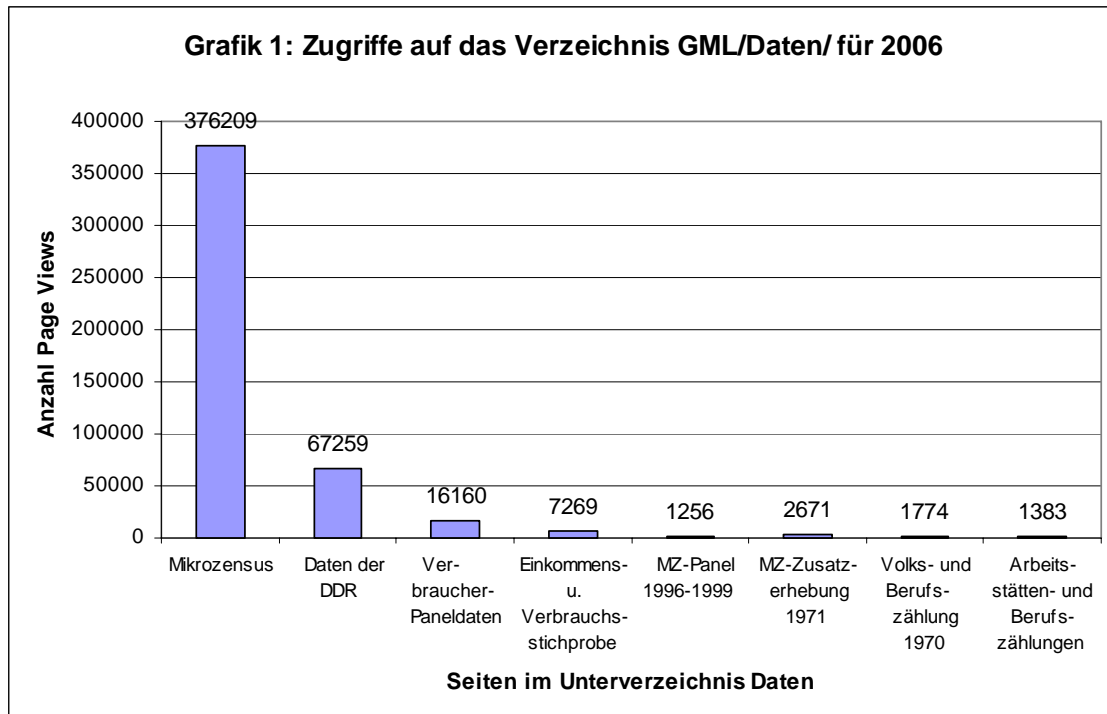
Tabelle 3: Webseiten und Zugriffe: /Daten

	Anzahl der Seiten	% von Gesamt (Tab.1)	PV	% von Ges. (Tab. 1)
2003	2.886	80.3	365.528	74,7
2004	3.606	81.7	541.393	75.6
2005	4.656	92.3	761.787	89.5
2006	5.170	89.7	475.349	86.3

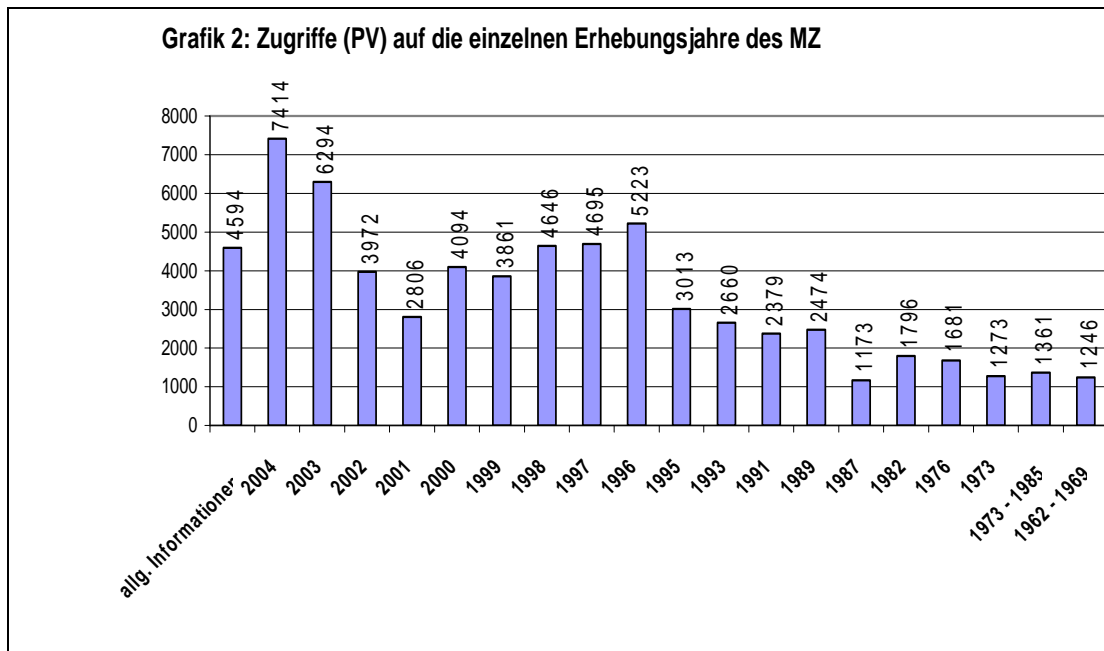
Dabei haben die Seiten mit Metadaten zu den verschiedenen Mikrozensus mit 81 Prozent den größten Anteil am Bereich Daten (Tabelle 4, Spalte 2). Ebenso entfallen 79 Prozent der Zugriffe im Bereich Daten auf den Mikrozensus (376.209 PV) (Tabelle 4, Spalte 4). Die restlichen Zugriffe verteilen sich auf die Informationsseiten zu den Daten der DDR (67.259 PV), Verbraucherpaneldaten (16.160 PV) und den Daten der Einkommens- und Verbrauchsstichprobe (7.269 PV) (Grafik 1). Selbst die Seiten mit Informationen zu den als „historisch“ einzuordnenden Daten der Volks- und Berufszählung (1.774 PV), der Mikrozensus-Zusatzerhebung 1971 (2.671 PV) und der Arbeitsstätten- und Berufszählungen von 1885-1970 (1.383 PV) werden häufig nachgefragt.

Tabelle 4: Webseiten und Zugriffe: /Mikrozensus

	Anzahl der Seiten	% von Daten (Tab.3)	PV	% von Daten (Tab.3)
2003	2.098	72.7	272.075	74,7
2004	2.809	77.9	416.130	76,9
2005	3.707	79.6	595.503	78,2
2006	4.201	81.3	376.209	79,1



Zur Beantwortung der Frage, wie sich die hohe Zahl an Page Views für die Mikrozensus auf die einzelnen Mikrozensusjahrgänge verteilt, dient Grafik 2. Am stärksten besucht wurden erwartungsgemäß die Seiten mit Informationen zu den neueren Jahrgängen der Mikrozensus Scientific Use Files 2004 und 2003 mit über 6.000 bzw. 7.000 Zugriffen. Zudem besteht ein großes Interesse an Informationen zu den Mikrozensus aus den 90er Jahren, wie aus Grafik 2 deutlich wird. Aber auch die älteren Mikrozensusjahrgänge sowie die absolut anonymisierten ZUMA-Files (1973-1985, 1962-1969) wurden mit jeweils über 1.000 Page Views im Jahr 2006 noch häufig aufgerufen, wobei bei den Zahlen für die Mikrozensus 1973, 1976, 1982 und auch 2004 zu berücksichtigen ist, dass diese Informationen nicht über das ganze Jahr verfügbar waren, sondern erst im Laufe des Jahres ins Web eingestellt wurden. Selbst im Verlauf des Jahres 2006 werden also in großem Umfang Informationen zu älteren und nicht nur den jeweils aktuellen Daten benötigt. Dies liegt u. E. darin begründet, dass bei vielen Forschungsprojekten nicht mehr nur mit dem jeweils aktuellen Datensatz gearbeitet wird, sondern mittlerweile aufgrund der zunehmenden Verfügbarkeit von Mikrozensus mehrere Zeitpunkte in die Untersuchungen einbezogen werden.

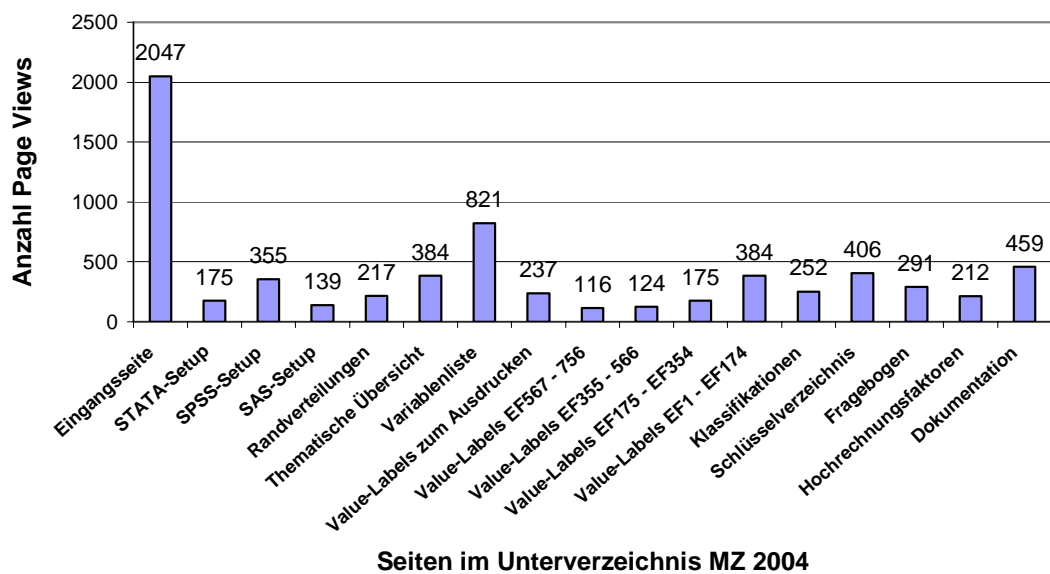
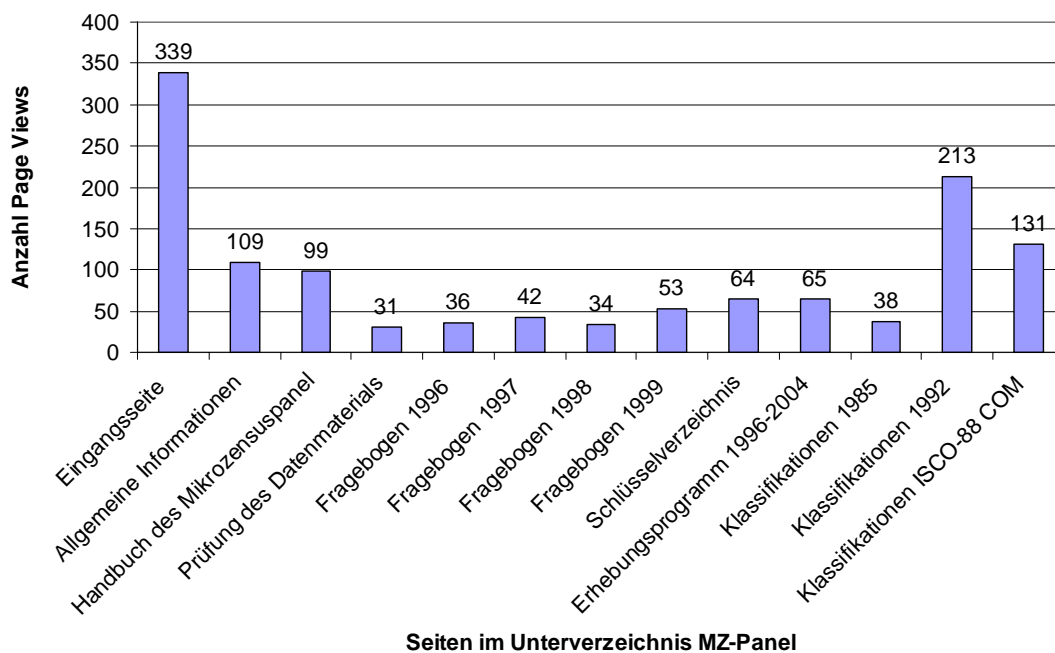


Hinter den Index-Seiten zu den einzelnen Mikrozensusjahrgängen, die in der Grafik 2 dargestellt sind, verbergen sich zahlreiche Detailinformationen zu den Scientific Use Files, wie amtliche und sozialwissenschaftliche Klassifikationen, Methodenberichte zur Aufbereitung, Schlüsselverzeichnisse etc. Aus Platzgründen wird hier exemplarisch nur der Zugriff auf die einzelnen Seiten der beiden aktuellsten Datenangebote, zum Mikrozensus 2004 (Grafik 3) sowie zum Mikrozensus-Panel (Grafik 4) dargestellt. Die Zahlen der Page Views zu den einzelnen Mikrozensen sind in Tabelle 5 dokumentiert.

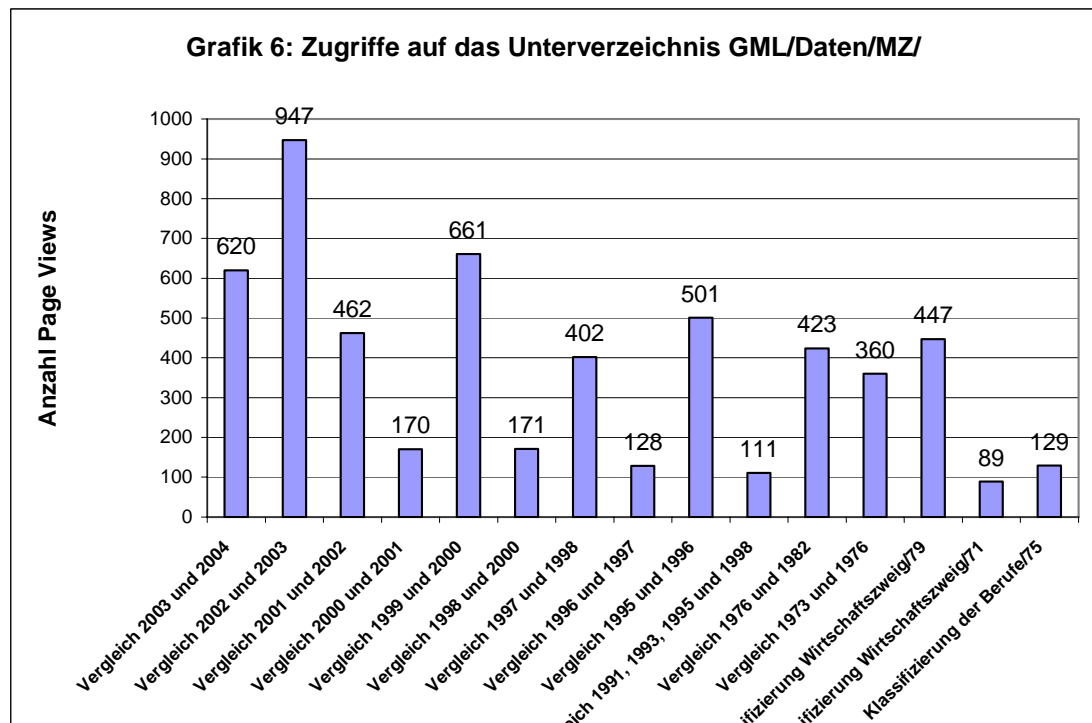
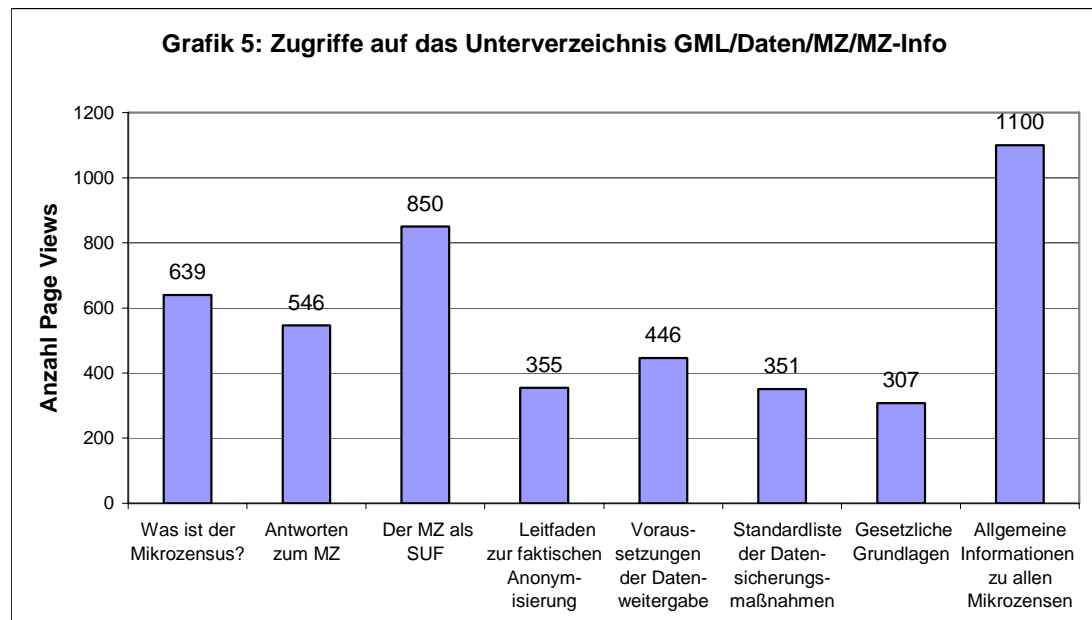
Die häufigsten Zugriffe erfolgten bei den meisten Mikrozensen auf die jeweilige Index-Seite, mit einem Durchschnitt von 19.5 Prozent sowie auf die Dokumentationen zur Datenaufbereitung (17.7 Prozent) (Tabelle 5). Beim MZ 2004 beträgt der Anteil der Zugriffe auf die Index-Seite 30.1 Prozent, gefolgt von den Variablenlisten (12.1 Prozent), Value-Labels (11.8 Prozent) und der Seite mit den Dokumenten zur Datenaufbereitung etc. (6.8 Prozent). Beim Mikrozensus-Panel ist es, abweichend vom sonstigen Muster, die Seite zu den Klassifikationen, wie z.B. ISCO oder „Klassifizierung des Berufs“ (30.2 Prozent), die am häufigsten aufgerufen wird, gefolgt von der Einstiegsseite (26.8 Prozent).

Tabelle 5: Page Views für alle Mikrozensus Jahrgänge 1962-2004, in Prozent eines Jahrgangs

Mikrozensus Jahr	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1993	1991	1989	1987	1982	1976	1973	Panel	73-85	62-69	Ges.
Eingangsseite	30,1	28,1	23,4	23,1	17,2	14,3	13,6	9,3	12,0	17,2	14,6	16,6	19,7	30,3	11,7	12,1	11,2	26,8	31,0	25,4	19,5
STATA-Setup	2,6	3,7	4,9	6,2	4,5	5,2	4,0	3,6	3,8	6,5	6,8	7,7	7,8	-	6,8	7,0	5,7	-	-	-	4,4
SPSS-Setup	5,2	5,2	5,2	8,2	4,6	6,8	4,4	3,7	3,6	5,9	8,2	6,7	8,6	-	6,4	7,2	6,1	-	-	-	5,1
SAS-Setup	2,0	3,0	4,5	5,7	4,5	4,8	3,8	3,2	3,2	5,2	5,7	6,3	6,6	-	3,5	3,8	3,9	-	-	-	3,7
Randverteilungen	3,2	3,5	4,4	6,3	4,7	5,6	4,5	3,9	3,1	5,8	6,9	7,9	7,3	-	5,3	5,8	5,1	2,5	-	-	4,4
Thematische Übersicht	5,7	8,8	6,8	6,0	4,1	4,2	3,9	2,5	3,0	4,6	5,3	6,1	5,8	-	6,5	6,8	6,2	-	-	-	4,8
Variablenliste	12,1	11,9	11,1	10,1	8,3	8,0	6,8	7,0	5,8	8,6	8,5	8,7	10,5	11,7	6,2	8,5	5,2	-	11,3	13,6	9,0
Value-Labels	11,8	2,8	3,2	4,3	2,5	3,1	2,8	2,7	2,2	5,1	5,1	5,8	4,7	22,2	5,3	4,7	3,8	-	17,9	24,3	5,6
Value-Labels zum Drucken	3,5	5,0	5,4	4,0	2,9	3,3	2,2	2,7	2,1	5,8	6,1	7,9	5,4	10,4	4,8	6,0	5,2	-	7,0	12,2	4,4
Klassifikationen	3,7	5,4	4,2	4,9	3,8	4,2	3,8	3,5	2,8	4,5	0,0	5,8	-	-	-	-	-	30,2	-	-	3,8
Schlüsselverzeichnis	6,0	4,4	4,1	4,5	3,3	3,8	4,1	1,9	2,5	3,2	7,4	9,4	7,4	-	10,6	9,0	12,5	5,1	-	-	4,6
Fragebogen	4,3	6,5	4,8	5,5	3,4	3,2	3,4	2,9	2,6	8,3	6,8	5,0	9,2	9,5	3,4	3,5	3,5	13,1	6,8	11,6	5,2
Abgrenzung/Hochrechn.	3,1	4,9	4,4	4,9	3,9	4,2	3,6	2,8	4,3	4,7	4,7	5,3	5,7	-	6,1	5,6	4,7	-	-	-	3,9
Dokumentation/Aufbereit.	6,8	6,9	13,5	6,2	32,5	29,3	39,0	50,4	37,0	10,3	0,9	0,8	1,3	-	16,1	12,1	14,4	-	-	-	17,7
Informationen zu Daten	-	-	-	-	-	-	-	-	-	-	-	-	-	15,9	7,3	7,9	6,4	8,6	11,8	13,0	1,5
Sonstiges	-	-	-	-	-	-	-	-	12,0	4,3	13,1	-	-	-	-	-	6,2	13,8	14,1	-	2,6
Gesamt N	6794	5347	3510	2636	3923	5199	4244	4567	4722	3013	2660	2379	2474	1173	1155	1103	1273	1264	1361	1246	58044

Grafik 3: Zugriffe auf das Unterverzeichnis GML/Daten/MZ 2004**Grafik 4: Zugriffe auf das Verzeichnis GML/Daten/MZ-Panel**

Neben den speziellen Informationen zu den einzelnen Scientific Use Files hat das GML auch zahlreiche allgemeine Informationen zum Mikrozensus erstellt, angefangen von den gesetzlichen Grundlagen der Mikrozensusbefragung bis zu den Bedingungen der Datenweitergabe an die Wissenschaft, die auf den Webseiten des GML dokumentiert werden (Grafik 5).



Diese Informationen wurden im Laufe des Jahres 2006 häufig aufgerufen, insbesondere die „Allgemeinen Informationen zum Mikrozensus“ (1.100 PV), die Informationen zum Mikrozensus als Scientific Use File (850 PV) oder die FAQ-Liste („Antworten zum MZ“, 546 PV). Neben diesen allgemeinen Informationen, die jahrgangsübergreifend sind, werden im Zuge der Aufbereitung auch spezielle Dokumentationen erstellt, die für die Nutzer der Daten wichtige Informationsquellen beim Arbeiten mit den Daten darstellen (Grafik 6). Eine solche „Standarddokumentation“ ist z.B. der Vergleich zwischen der jeweils aktuellen Erhebung des Mikrozensus mit der vorangegangenen Erhebung, in der die Änderungen zwischen den Erhebungen detailliert beschrieben werden. Diese Seiten wurden von den Nutzern regelmäßig nachgefragt; die häufigsten Page Views erzielte dabei der Vergleich zwischen den Erhebungen 2002 und 2003 (947 PV).

2.3 Service

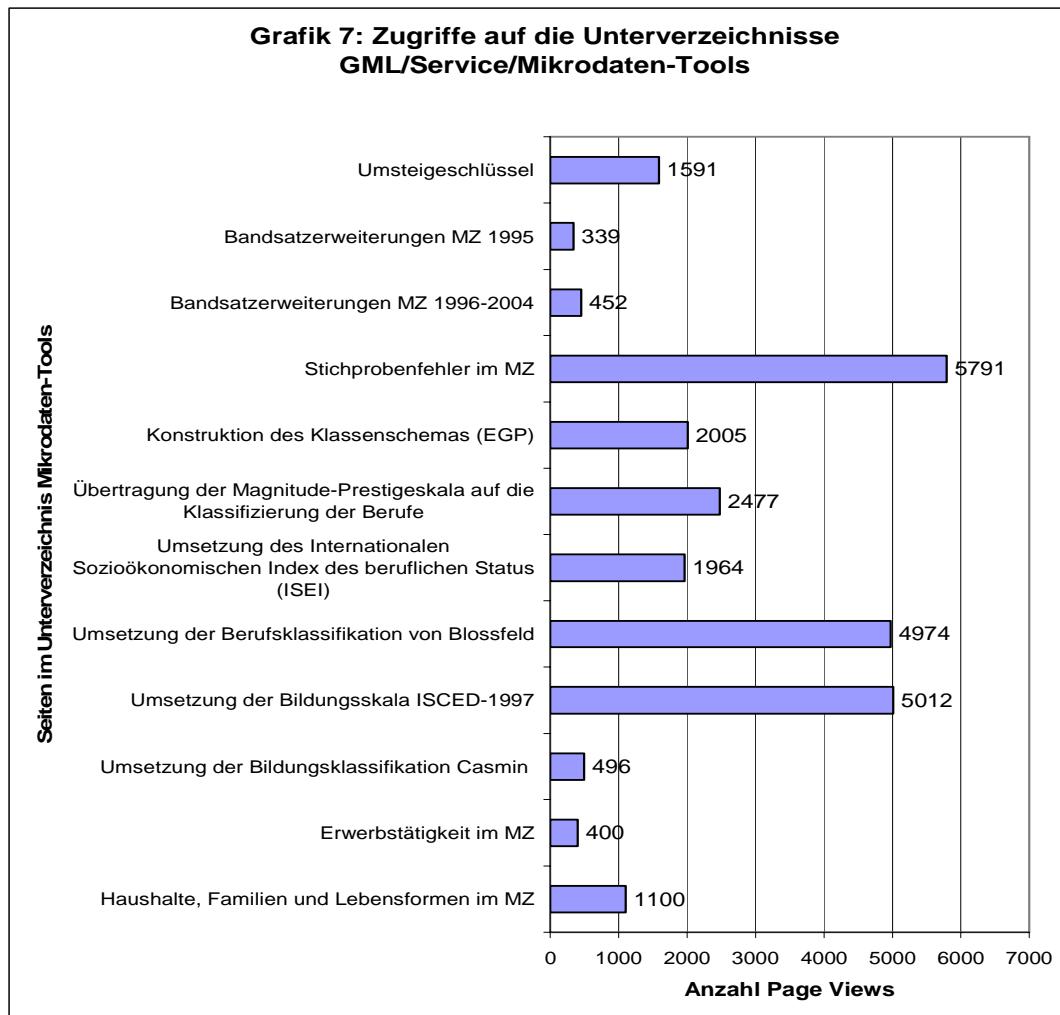
Ein zentrales Serviceangebot sind die so genannten Mikrodaten-Tools (Tabelle 6). Hierzu zählt die Umsetzung verschiedener sozialwissenschaftlicher Instrumente mit den Daten des Mikrozensus, wie z.B. Bildungs- und Berufsklassifikationen, die Erläuterung von Konzepten, die in den Daten der amtlichen Statistik angewandt werden (z.B. Erwerbs- oder Unterhaltskonzept) oder methodische Hinweise zum Arbeiten mit den Daten. Das Angebot dazu wurde im Jahr 2006 um mehrere Tools erweitert, was sich in einer Erhöhung der Seitenzahlen ausdrückt. Im Schnitt wurden die Seiten zu den Mikrodaten-Tools pro Tag annähernd 50-mal abgerufen.

Tabelle 6: Webseiten und -Zugriffe: /Mikrodaten-Tools

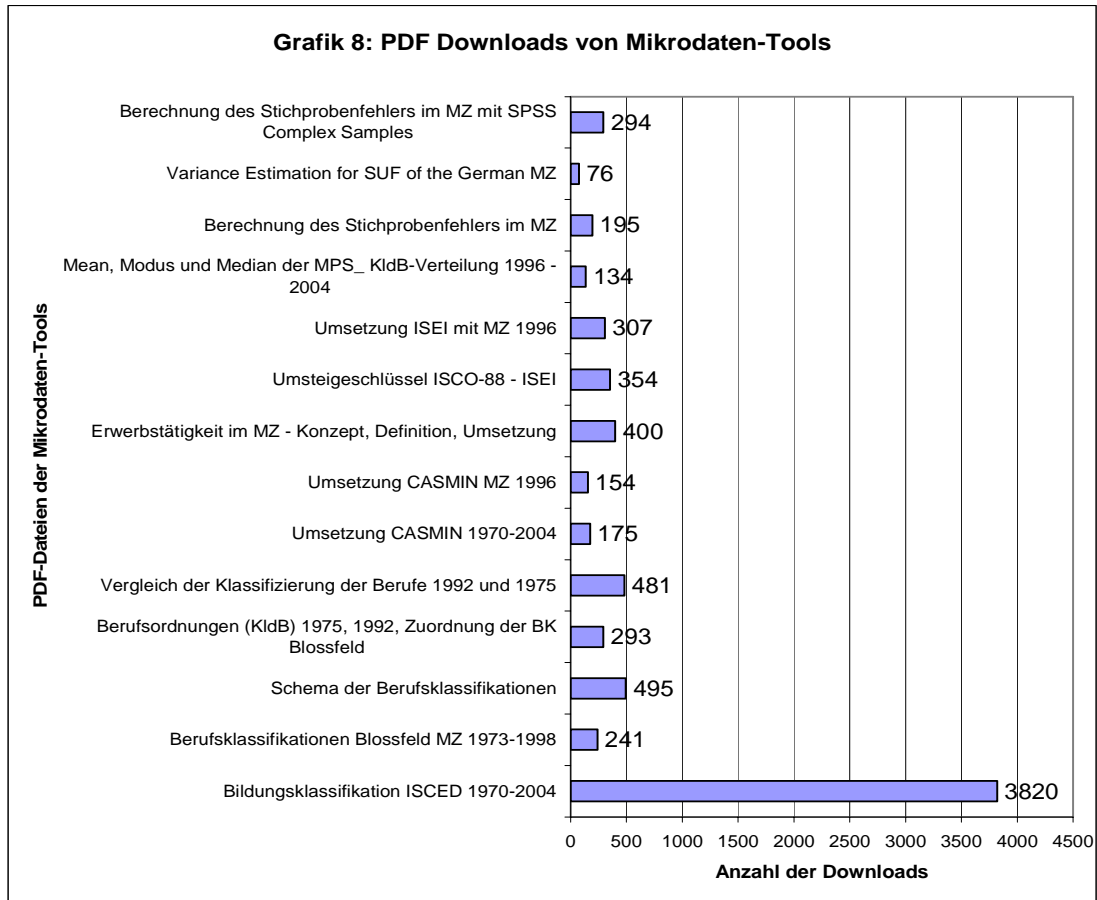
	Anzahl der Seiten	PV	PV pro Tag
2003	106	14.537	39
2004	114	21.053	59
2005	127	19.206	53
2006	258	17.548	48

Wie aus Grafik 7 zu entnehmen ist, wurde im Jahr 2006 vor allem auf die Seite zur Berechnung des Stichprobenfehlers im Mikrozensus (5.791 PV) und die beiden Tools zu den Umsetzungen der Berufsklassifikation von Blossfeld (4.974 PV) sowie der Bildungsklassifikation ISCED (5.012 PV) zugegriffen.

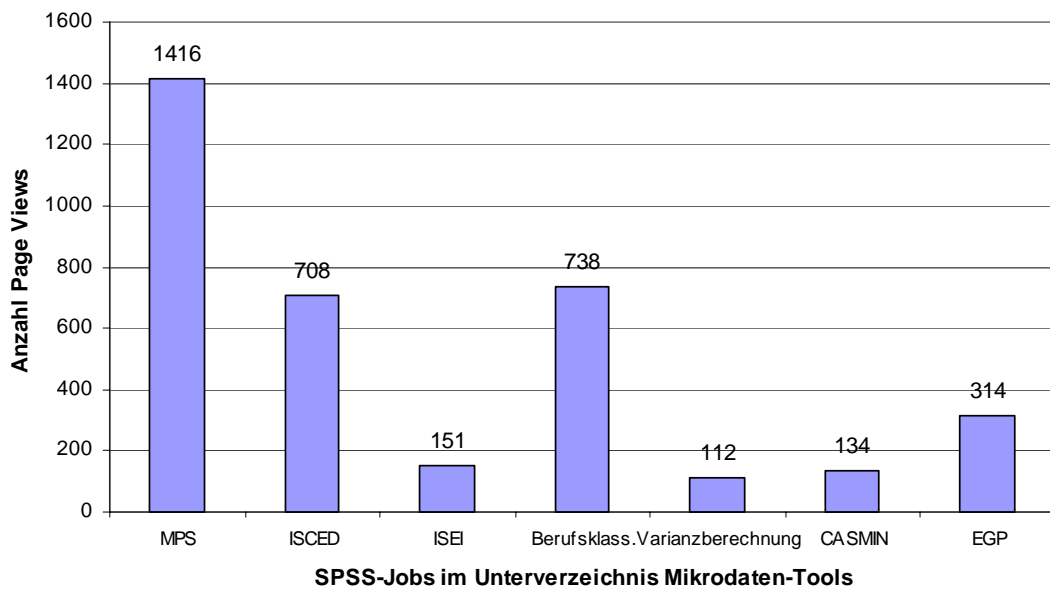
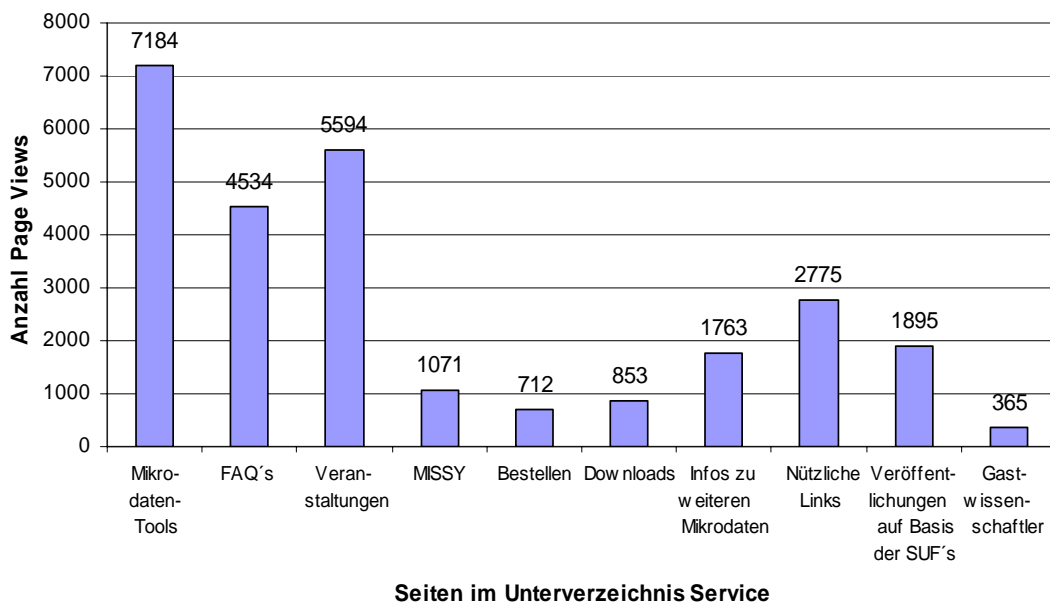
Die Page Views zu den Unterverzeichnissen im Folder Mikrodaten-Tools enthalten dabei sowohl Html-Seiten als auch PDF-Dateien.



Grafik 8 enthält nur die Downloads von PDF-Dateien in den Verzeichnissen der Mikrodaten Tools. Bei den PDF-Dateien handelt es sich in der Regel um Arbeits- oder Methodenberichte, in denen die Entwicklung der Tools beschrieben wird. Hier wurde insbesondere der Bericht zur Erstellung der ISCED Klassifikation mit 3.820 Downloads stark nachgefragt, aber auch die Downloadzahlen anderer Berichte, die meist zwischen 300 und 500 liegen, zeigen, dass das Angebot dieser „Werkzeuge“ bei dem Nutzer auf ein hohes Interesse stößt. Während die Berichte die Entwicklung von Tools beschreiben, werden die Tools selbst in der Regel als SPSS Syntax-File und neuerdings ebenfalls als Stata do-file zum Download angeboten.



Grafik 9 enthält daher die Zahlen der Downloads von SPSS-Jobs, in denen für das jeweilige sozialwissenschaftliche Instrument, die konkrete Syntax zur Erzeugung von Klassifikationen etc. mit den Daten des Mikrozensus Scientific Use Files angeboten wird. Am häufigsten wurden von den Nutzern die Routinen zur Erzeugung der Magnitude Prestige Skala (1.416) und der Berufsklassifikation von Blossfeld (738) sowie der ISCED Bildungsklassifikation (708) von den Webseiten des GML heruntergeladen.

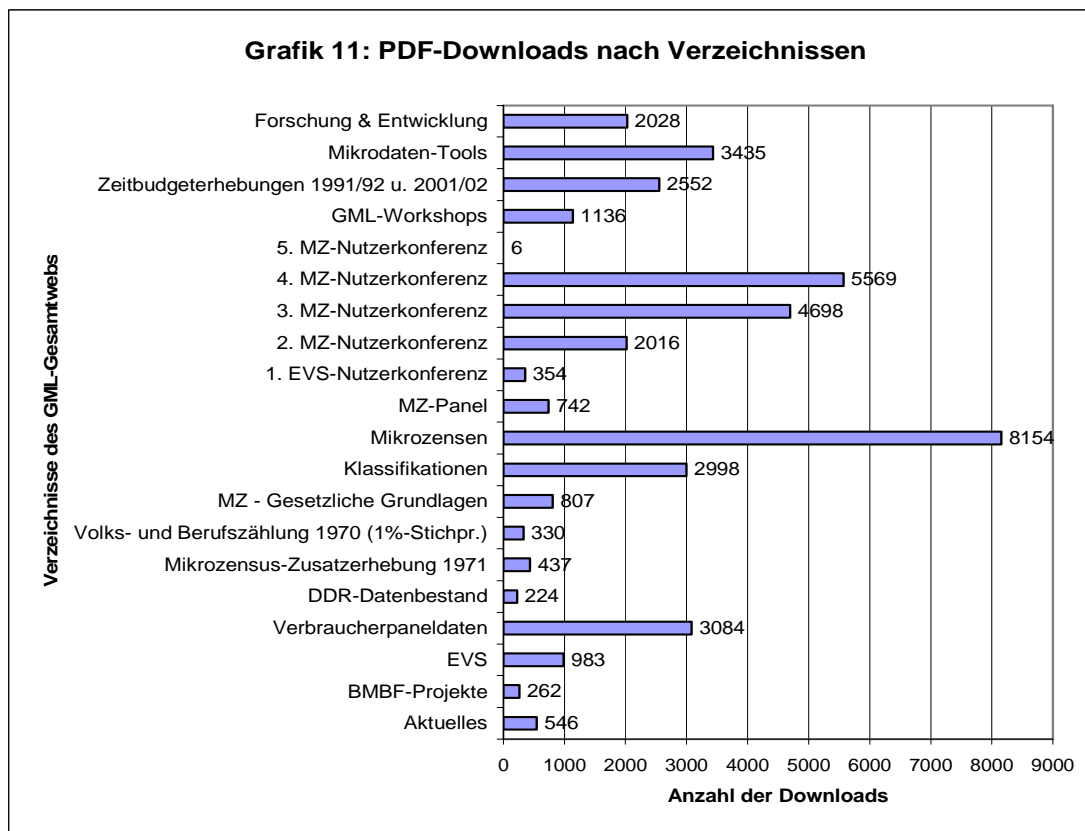
Grafik 9: Downloads von Tools (SPSS-Jobs)**Grafik 10: Zugriffe auf das Unterverzeichnis GML/Service**

In der Rubrik „Service“ befinden sich auch die Seiten mit den Ankündigungen zu den Workshops und Nutzerkonferenzen des GML, zu sozialwissenschaftlichen Konzepten und Klassifikationen (Tools), zur FAQ-Liste etc. Am häufigsten aufgerufen wurden Seiten im Bereich der Mikrodamentools (7.184 PV), im Bereich zu den Veranstaltungen (5.594 PV) sowie die FAQ-Liste zum Mikrozensus (4.534 PV), in der häufig gestellte Fragen – und deren Beantwortung – zum Mikrozensus und zum Scientific Use File dokumentiert werden.

Grafik 11 enthält die Anzahl der Downloads von PDF-Dateien der einzelnen Verzeichnisse des GML Gesamtwebs. Wie auch schon bei der Grafik 8 vermittelt die Anzahl der Downloads von PDF-Files ein realistischeres Bild der Nutzung, als es die Anzahl der Page Views unter Umständen auszudrücken vermag.

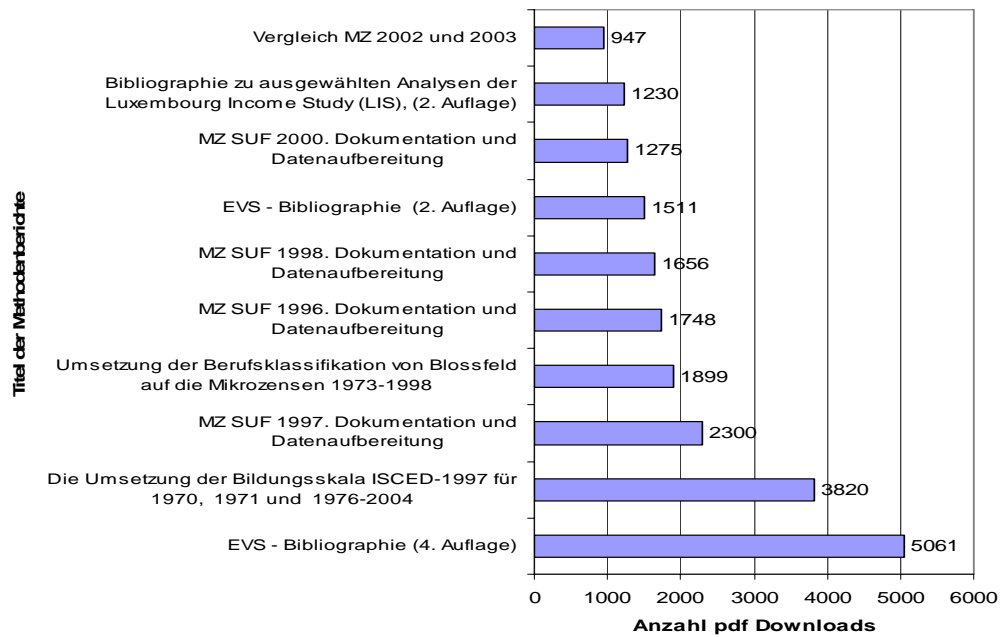
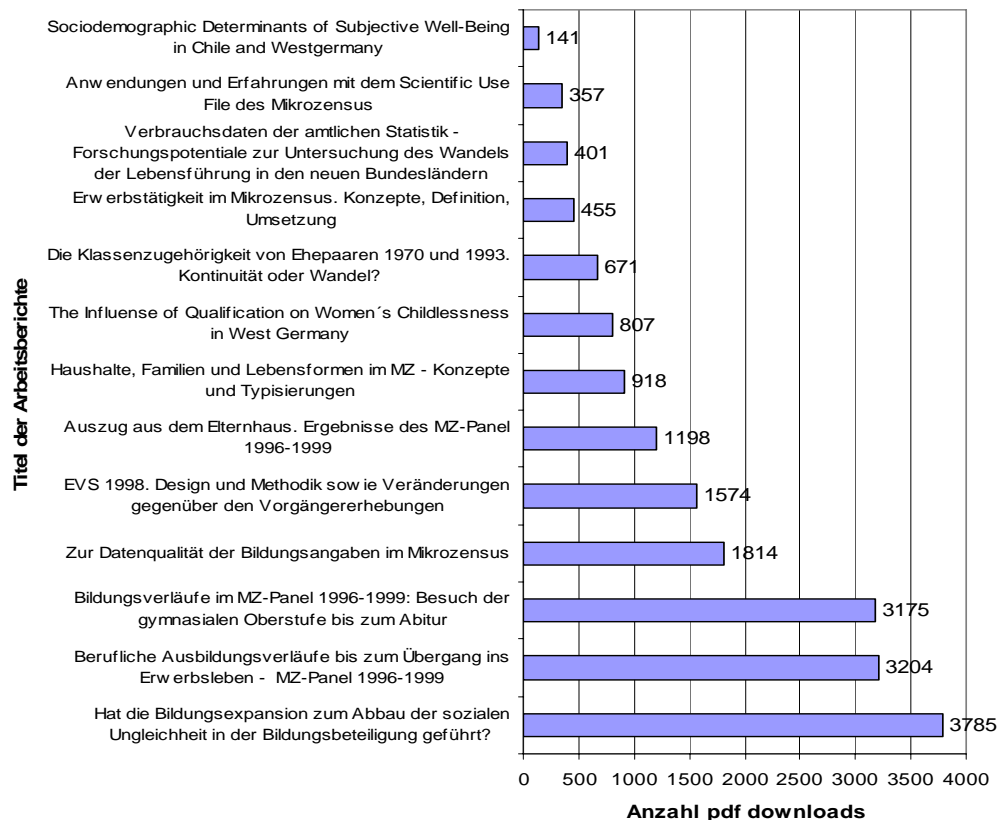
Sehr stark nachgefragt werden einerseits Dokumentationen und Informationen, die sich auf die Arbeit mit dem Mikrozensus beziehen, also z.B. Klassifikationen wie die Klassifizierung der Berufe, ISCO Code etc. (2.998 Downloads), die Dokumente zu den Variablen und Value Labels der einzelnen Mikrozensen (8.154) und die verschiedenen Mikrodaten-Tools (3.435).

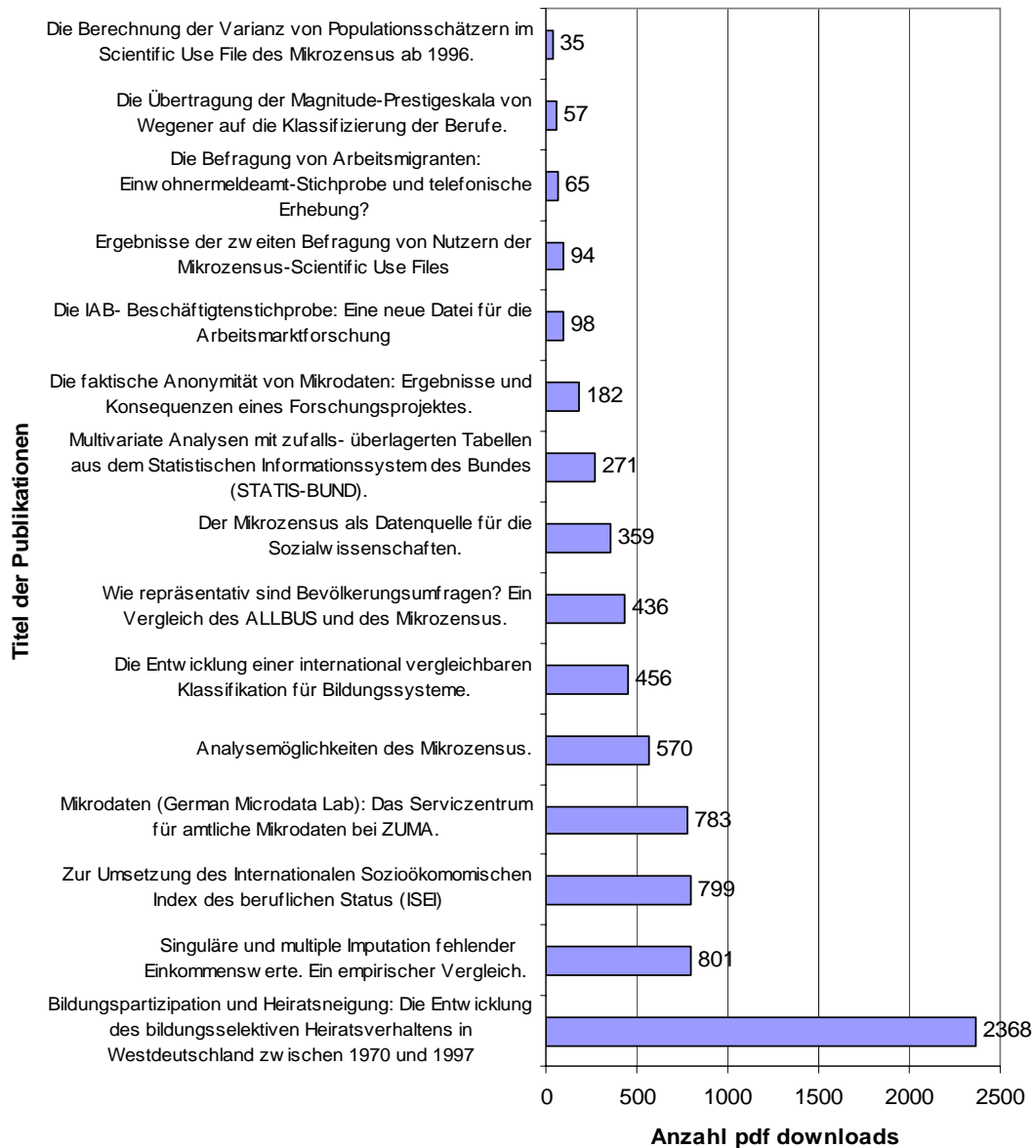
Andererseits fanden auch inhaltliche Beiträge und Informationen ein starkes Interesse. So verbergen sich hinter den PDF-Downloads der einzelnen Nutzerkonferenzen sowohl die Abstracts der Beiträge, in wesentlich höherer Zahl aber die vollständigen Referate und Publikationen. Diese hohen Downloadzahlen zeigen, dass die Beiträge – auch die der bereits älteren Nutzerkonferenzen – sehr stark nachgefragt werden.



2.4 Publikationen

Zum Abschluss betrachten wir die Nutzung der verschiedenen Publikationen der Mitarbeiter des German Microdata Lab. Dabei handelt es sich allerdings nur um die Publikationen, die in der Reihe der GESIS-ZUMA Arbeits- bzw. Methodenberichte oder z.B. in den ZUMA-Nachrichten erschienen sind, also um PDF-Dateien, die auf dem GESIS-Server zur Verfügung gestellt werden. Die über sonstige Zeitschriften erschienenen Publikationen werden somit nicht erfasst. Unter den Top Ten der Methodenberichte (Grafik 12) befinden sich fünf Dokumentationen zu Aufbereitungen verschiedener Scientific Use Files, drei Berichte mit Bibliographien sowie zwei Berichte zur Konstruktion von Tools. An der Spitze der Downloads stehen dabei die Bibliographie zu Publikationen mit der Einkommens- und Verbrauchsstichprobe (5.061) sowie der Bericht zur Konstruktion der ISCED-Skala (3.820). Bei den eher inhaltlich orientierten Arbeitsberichten (Grafik 13 und 14) sind es vor allem die bildungsspezifischen Arbeitspapiere sowie die Arbeitspapiere, die sich mit dem Mikrozensus-Panel beschäftigen, die am stärksten nachgefragt werden.

Grafik 12: Top Ten der Methodenberichte**Grafik 13: Zugriffe auf Arbeitsberichte des GML**

Grafik 14: Zugriffe auf sonstige Publikationen

3. Zusammenfassung

Die Auswertung der Web-Statistik der vom GML angebotenen Seiten zeigt die häufige Nutzung dieses Services. Insbesondere die Metainformationen, aber auch Tools zu den Daten oder wissenschaftliche Publikationen werden intensiv von den Nutzern nachgefragt.

Seitdem der Mikrozensus als Scientific Use File kostengünstig verfügbar ist, haben circa 170 universitäre und außeruniversitäre Forschungseinrichtungen Daten des Mikrozensus angefordert. Darunter befinden sich alle in Frage kommenden empirisch arbeitenden sozial- und wirtschaftswissenschaftlichen Forschungseinrichtungen (DIW, ZEW, WZB, MPI etc.) und viele Lehrstühle an Universitäten. In der Regel arbeiten zudem mehrere Forscher an den Instituten mit den Daten, so dass die Zahl der Wissenschaftler, die mit den Scientific Use Files arbeiten, deutlich höher liegt, als es die Zahl der Einrichtungen abzubilden vermag. Die hohen Nutzungszahlen der Downloads oder Page Views zeigen somit den hohen Gebrauchswert sowie die Nützlichkeit der vom GML bereit gestellten Dienstleistungen für die Forscher an.

Ebenso sind diese Nutzungszahlen ein Indiz dafür, dass das Mikrodateninformationssystem MISSY, das bislang als Prototyp vorliegt und in das die Metadaten zu allen verfügbaren Scientific Use Files der Mikrozensen eingebunden werden sollen, für den Nutzer einen wesentlichen Fortschritt für die effiziente Arbeit mit den Daten darstellen wird, da dann die bislang über die GML Webseiten vorliegenden Einzelinformationen integriert und funktional verknüpft für den Nutzer angeboten werden.